

(19) KOREAN INTELLECTUAL PROPERTY OFFICE

KOREAN PATENT ABSTRACTS

(11)Publication number: 1020020043115 A
(43)Date of publication of application: 08.06.2002

(21)Application number: 1020000072590

(71)Applicant: IPENTEC. INC.

(22)Date of filing: 01.12.2000

(72)Inventor: KANG, HAN JONG

(30)Priority:

(51)Int. Cl G06F 17/00

(54) TEXT TAGGING METHOD AND RECORDING MEDIUM

(57) Abstract:

PURPOSE: A text tagging method and a recording medium are provided to apply XML structure information on a text by reading a general document to be converted into an XML document as a text form on one program. CONSTITUTION: If a user applies to structure information and designates a general text document to be converted into an XML document or a non-converted XML document, an XML editor reads a corresponding file and displays the structure information on an XML tree window and displays a text file contents on an XML editor window file. If the user selects a text area(S50), a selection signal corresponded to the text area is inputted to a text tagging manager. If an element which is inserted into a parent node of the selected area(for example technique research institute) is existed(S52), the text tagging manager suggests a candidate element(for example employee) which is tagging on a text tagging candidate window(S54). The text tagging manager judges whether an element tagging selection or an element lower selection is existed from the user(S56). If an element tagging is selected, a selection signal corresponded to the selection is inputted to the text tagging manager, and the text tagging manager inserts the selected area to contents of the selected candidate element(S58).

copyright KIPO 2003

Legal Status

Date of request for an examination (20001201)

Notification date of refusal decision (00000000)

Final disposal of an application (registration)

Date of final disposal of an application (20040319)

Patent registration number (1004374470000)

Date of registration (20040615)

Number of opposition against the grant of a patent ()
Date of opposition against the grant of a patent (00000000)
Number of trial against decision to refuse ()
Date of requesting trial against decision to refuse ()
Date of extinction of right ()

공개특허 등2002- 0043115

(19) 대한민국특허청(KR)
 (12) 공개특허공보(A)

(51) Int. Cl. 7
 G06F 17/00

(11) 공개번호 등2002- 0043115
 (43) 공개일자 2002년06월08일

(21) 출원번호 10- 2000- 0072590
 (22) 출원일자 2000년12월01일

(71) 출원인 (주)아이펜텍
 강명필
 서울특별시 강남구 역삼동 651- 3 한신빌딩 3층

(72) 발명자 강한종
 서울특별시성동구행당동147삼부아파트106동1101호

(74) 대리인 박래봉

심사청구 : 있음

(54) 텍스트 태깅 방법 및 기록매체

요약

본 발명은 하나의 프로그램상에 XML 문서로 변환하고자 하는 일반 문서를 텍스트 형태로 읽어 들여 놓고 이 텍스트상에 XML 구조 정보를 적용시키도록 한 텍스트 태깅 방법 및 기록매체에 관한 것으로, 변환하기 위한 비 XML 텍스트 및 변환 과정에서 생성된 XML 텍스트를 하나의 문서로 통합하여 관리하고, 그 XML 텍스트 및 비 XML 텍스트의 선택된 영역에 대해 XML 문서 규칙에 따른 적용가능한 태깅정보를 제공하여, 읽어들인 XML 텍스트 문서 또는 비 XML 텍스트 문서를 대상으로 상황에 맞는 후보 엘리먼트를 제시하여 선택되는 후보 엘리먼트의 내용으로 태그를 삽입시킴으로써, XML 텍스트 문서로의 변환시 간편하게 태깅할 수 있게 해 주고, 그로 인해 텍스트 내용에 구조 정보를 적용시키는 과정 및 시간이 단축된다.

대표도
 도 10

색인어
 XML 텍스트 문서, 문서 변환, 태깅, 후보 엘리먼트, 비 XML 텍스트 문서

명세서

도면의 간단한 설명

도 1은 본 발명의 방법이 적용되는 시스템 구성도,

도 2는 본 발명의 방법이 적용되는 시스템의 작업 원도우의 예,

도 3은 본 발명의 방법이 적용된 시스템에서 문서 열기를 하여 문서 편집과정으로 진입하기까지의 동작을 설명하는 흐름도,

도 4는 도 3에서 사용자가 XML 파일 열기를 한 경우의 화면 예,

도 5는 도 3에서 사용자가 텍스트 파일을 열고 XML 파일을 템플릿으로 사용한 경우의 화면 예,

도 6은 도 3에서 사용자가 텍스트 파일을 열 때 환경설정(문서 형태 정의 선택)을 하는 화면 예,

도 7은 도 3에서 사용자가 텍스트 파일을 열고 XML 파일을 템플릿으로 사용하지 않은 경우의 화면 예,

도 8은 본 발명의 방법이 적용된 시스템에서 XML 문서 열기에서의 오류 처리과정을 설명하는 흐름도,

도 9는 도 8에서 오류 발생시의 화면 예,

도 10은 본 발명의 실시예에서 텍스트 태깅 후보 엘리먼트를 제시하는 과정을 설명하는 흐름도,

도 11 내지 도 13은 도 10의 텍스트 태깅 후보 엘리먼트 제시 과정 설명에 채용되는 화면 예들이다.

※ 도면의 주요부분에 대한 부호의 설명

10 : 문서 객체 모델(DOM) 12 : 문서 트리 관리자

14 : 파서 16 : XML 편집기

18 : 후보 엘리먼트 관리자 20 : 텍스트 태깅 관리자

22 : 속성 관리자 24 : 오류 관리자

30 : XML 트리 원도우 32 : XML 에디터 원도우

34 : 엘리먼트 후보 목록 원도우 36 : 속성 목록 원도우

38 : 텍스트 태깅 후보 원도우 40 : 에러 리스트 원도우

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 텍스트 태깅 방법 및 기록매체에 관한 것으로, 보다 상세하게는 비 XML 텍스트 문서에 구조 정보를 적용시켜 XML 문서로 변환시킬 때 선택된 텍스트 영역에 적절한 태그를 적용시키는 방법 및 그 방법이 프로그램화되어 기록된 기록매체에 관한 것이다.

종래의 문서는 다양한 문서편집기를 사용하여 다양한 형식의 문서로 생성되었으며 저자의 기본 스타일에 따라 작성되므로 구조화되어 있지 않아 문서의 교환을 통한 공유를 위해서는 별도의 변환 작업이 필요하다.

문서의 개념도 기존의 문자 중심에서 음성, 동영상 등의 멀티미디어가 포함되는 멀티미디어 문서라는 개념으로 바뀌고 있다.

한편, 정보 사회가 성숙됨에 따라 시스템 및 응용에 독립적인 문서 정보의 처리 요구가 커지게 되었고 이에 따라 문서 정보의 교환, 검색 등의 처리를 위한 문서 구조 표준의 필요성이 대두되어 미국표준협회나 ISO 등에서는 SGML 등 멀티미디어 문서 교환 형식에 대한 국제 표준을 발표하고 있다.

상기 SGML(Standard Generalized Markup Language)은 한 문서에서 논리 구조와 내용 구조를 기술하기 위한 메타언어로서, 다양한 응용들 사이에 구조화된 데이터를 상호 교환하기 위한 도구이다.

SGML로 문서를 작성할 때 문자의 형태나 줄 간격 같은 실제 문서를 꾸미는 것이 아니라 태그(tag)라는 것을 사용하여 내용과 형식의 구분없이 작성한다. 즉, 주제 또는 표 등을 구분하지 않고 동일한 방식인 태그를 이용하여 표시하는 것이다. 우리가 지금 웹에서 쓰는 HTML문서도 SGML문서의 한 종류라고 볼 수 있다. SGML문서는 태그에 의해서 작성되기 때문에 SGML 사용자 그룹에서 여러 종류의 태그를 정의하고 있다.

SGML문서는 문서를 기술하기 위해 사용된 구문을 설명하는 SGML 선언(SGML declaration), 엔티티(entity), 요소(element), 속성(attribute) 선언을 통해 문서의 구조와 내용을 정의하는 문서 형태 정의(DTD : Document Type Definition)와 실제 DTD에 따른 내용과 태그로 구성된 실제 문서인 문서 인스턴스(document instance)로 이루어진다.

HTML(Hyper Text Markup Language)은 SGML언어의 한 응용으로써, 웹에서 하이퍼 텍스트 문서를 만들기 위하여 쓰이는 간단한 마크업 언어이다. 마크업 언어는 문서의 중간에 문자나 문장 형태의 명령어를 삽입하는 방식의 문서 표현 언어로서 명령어는 태그를 사용하여 삽입한다.

HTML에서는 태그만으로 정보를 원하는 대로 표현할 수 있기 때문에 별도의 프로그래밍 지식이 없어도 인터넷상에서 자신이 가지고 있는 정보를 쉽게 제공할 수 있으며, 그라피클 기반으로 하여 정보가 표시되기 때문에 사용하기가 쉽다.

상기 SGML이 용통성과 재사용성이 뛰어나며 시스템이나 플랫폼에 독립적인 개방형 표준이라는 장점에도 불구하고, 복잡한 SGML을 완전히 표현할 수 있는 소프트웨어의 개발이 어렵다는 문제가 있다.

그리고, 작성하기 쉽기 때문에 급속히 확산될 수 있었던 HTML은 원래의 목적이 브라우저를 통한 문서의 표현에 있으므로 HTML 문서 정보의 재활용에는 어려움이 있다.

따라서, 방대한 양의 HTML 파일을 재구성하려면 다시 작성해야 하는 문제가 있으므로 브라우저가 아닌 웹에서의 SGML 사용의 필요성이 대두되었다.

이와 같은 HTML의 장점과 SGML의 장점을 혼합한 새로운 표준이 XML이다.

XML(eXtensible Markup Language)은 SGML의 자주 쓰이지 않는 부분을 대폭 축소하였고, 링크나 문서 출력 부분은 현재의 HTML 기능을 포함한다. 즉, XML은 SGML의 복잡한 옵션을 제거한 SGML의 서브셋이며, HTML보다 복잡한 문서 생성이 가능하도록 확장한 형태로서, 웹상에서 구조화된 문서를 전송할 수 있도록 표준화된 문자 형식의 메타언어라고 할 수 있다.

이와 같이 종래의 문서작성 도구(워드프로세서 등의 문서 편집기)들은 거의 대부분이 자체 형식과 순수 텍스트 형식을 동시에 지원한다. 그러므로 모든 문서들을 텍스트 형식의 문서로 변환한 후에, 이 텍스트 문서를 대상으로 XML 문서로의 변환을 꾀하는 방법이 범용적인 XML 문서 변환을 달성하는 방법이라고 할 수 있다.

그러나, 텍스트 문서는 아무런 논리적 구조를 가지지 않으며, 가장 단순한 플랫 문서이다. 이러한 텍스트 문서를 사람의 손을 거치지 않고 구조화된 XML 문서로 변환하기는 어렵다. 내용상으로 특정 공통 규칙을 가지고 작성된 텍스트 문서들에 대해서는 비록 플랫 문서일지라도 규칙을 적용시켜 XML 문서로 변환을 꾀할 수 있겠지만, 일반적인 모든 텍스트

문서에는 적용될 규칙을 만들어 적용시킬 수 없다. 그러므로 텍스트 문서를 구조화된 XML 문서로 변환하기 위해서는 사람이 대상 문서에 구조 정보를 적용시켜 변환할 수 밖에 없다. 얼마나 편리하고 쉽게 그리고 효율적으로 빠르게 텍스트 문서에 구조 정보를 적용시켜 변환을 할 수 있는가가 가장 중요한 요소가 된다.

종래의 XML 문서 편집기의 경우, 현재 작업중인 문서의 구조 정보를 트리(tree) 형태로 화면상에 표현하는 기능, 현재 작업중인 문서를 텍스트(text) 형태로 화면상에 표현하는 기능, 현재 작업중인 문서상의 위치 또는 사용자 선택한 노드에 따라서 삽입이 가능한 엘리먼트 후보들을 화면상에 제시해 주는 기능, 문서상의 현재 위치하고 있는 엘리먼트 또는 선택한 엘리먼트가 속성을 가지고 있을 경우에 설정 가능한 속성 목록과 속성값을 화면상에 보여주며 그 값을 설정하거나 변경할 수 있도록 해주는 기능 등을 가지고 있다.

이러한 기능들을 갖춘 종래의 XML 문서 편집기는, 신규로 XML 문서를 생성하거나 기 작성된 XML 문서를 수정/보완하는데는 적합하다.

그러나, 이미 일반 문서 편집기(예컨대, 워드 프로세서, 아래한글, MS Word 등)를 사용하여 생성시킨 많은 문서들을 XML 문서로 변환하는 것이 요구되고 있는 까닭에 종래의 XML 문서 편집기로는 기 생성된 일반 문서들을 XML 문서로 변환하는 것이 부적합하다.

예를 들면, 종래의 XML 문서 편집기로 변환작업을 수행하기 위해서는 일반 문서 편집기와 XML 문서 편집기를 동시에 병행하여 실행시켜야 한다. 이러한 상태에서, 먼저 XML 문서 편집기에서 문서의 논리 구조를 생성시킨 후에 일반 문서 편집기에서 대응되는 텍스트를 선택하여 XML 문서 편집기상의 대응되는 논리 구조 영역안에 붙여 넣기를 해야 되는 번거러운 작업을 반복해야만 한다.

그에 따라, 일반 문서를 XML 문서로 변환하기 위해서는 작업자가 2개의 실행 프로그램(즉, 일반 문서 편집기 및 XML 문서 편집기의 실행 프로그램)을 운용하여 그 프로그램간을 번갈아 가며 XML 문서를 완성시켜 나간다. 이를 위해서는 작업자의 주의깊은 세심한 작업이 필요하므로 오류 발생율이 많은 실정이다.

발명이 이루고자 하는 기술적 과제

본 발명은 상기한 종래의 사정을 감안하여 안출된 것으로, 하나의 프로그램상에 XML 문서로 변환하고자 하는 일반 문서를 텍스트 형태로 읽어 들여 놓고 이 텍스트상에 XML 구조 정보를 적용시키도록 한 텍스트 태깅 방법을 제공함에 그 목적이 있다.

본 발명의 다른 목적은 하나의 프로그램상에 XML 문서로 변환하고자 하는 일반 문서를 텍스트 형태로 읽어 들여 놓고 이 텍스트상에 XML 구조 정보를 적용시키도록 한 텍스트 태깅 프로그램이 저장된 기록매체를 제공함에 있다.

발명의 구성 및 작용

상기와 같은 목적을 달성하기 위하여 본 발명의 바람직한 실시예에 따른 텍스트 태깅 방법은, 변환하기 위한 비 XML 텍스트 및 변환 과정에서 생성된 XML 텍스트를 하나의 문서로 통합하여 관리하는 제 1과정; 및 그 XML 텍스트 및 비 XML 텍스트의 선택된 영역에 대해 XML 문서 규칙에 따른 적용가능한 태깅정보를 제공하여 논리 구조를 적용시키는 제 2과정을 포함한다.

본 발명의 다른 실시예에 따른 텍스트 태깅 방법은, 편집 영역에 나타나는 문서내의 XML 텍스트 영역 및 비 XML 텍스트 영역중에서 일부 영역이 선택되면 상기 선택된 일부 영역에 태깅가능한 후보 엘리먼트의 존재여부를 파악하여 화면상에 표시하는 제 1과정; 및 그 화면표시된 후보 엘리먼트에 대한 태깅이 선택되면 그 선택된 일부 영역에 상기 선택된 엘리먼트의 내용으로 태그를 삽입하고, 상기 표시된 후보 엘리먼트에 대한 하부 엘리먼트 선택이 있으면 상기 선택된

후보 엘리먼트의 하부 엘리먼트를 조사하여 화면상에 표시하여 태깅 선택을 대기하는 제 2과정을 포함한다.

그리고, 본 발명의 실시예에 따른 기록매체는, 변환하기 위한 비 XML 텍스트 및 변환 과정에서 생성된 XML 텍스트를 하나의 문서로 통합하여 관리하는 제 1과정; 및 그 XML 텍스트 및 비 XML 텍스트의 선택된 영역에 대해 XML 문서 규칙에 따른 적용가능한 태깅정보를 제공하여 논리 구조를 적용시키는 제 2과정의 순서로 진행되는 텍스트 태깅 프로그램이 저장된다.

본 발명의 다른 실시예에 따른 기록매체는, 편집 영역에 나타나는 문서내의 XML 텍스트 영역 및 비 XML 텍스트 영역 중에서 일부 영역이 선택되면 상기 선택된 일부 영역에 태깅가능한 후보 엘리먼트의 존재여부를 파악하여 화면상에 표시하는 제 1과정; 및 그 화면표시된 후보 엘리먼트에 대한 태깅이 선택되면 상기 선택된 일부 영역에 상기 선택된 엘리먼트의 내용으로 태그를 삽입하고, 상기 표시된 후보 엘리먼트에 대한 하부 엘리먼트 선택이 있으면 상기 선택된 후보 엘리먼트의 하부 엘리먼트를 조사하여 화면상에 표시하여 태깅 선택을 대기하는 제 2과정의 순서로 진행되는 텍스트 태깅 프로그램이 저장된다.

이하, 본 발명의 실시예에 따른 텍스트 태깅 방법에 대하여 첨부된 도면을 참조하여 설명하면 다음과 같다.

도 1은 본 발명의 방법이 적용되는 시스템 구성도로서, 처리 응용이나 스크립트로부터 XML 텍스트 문서의 요소와 속성 등을 다루는 향상된 문서 객체 모델(10)(Document Object Model : 이하 DOM+ 라고 함); 상기 DOM+(10)과 연동하여 XML 트리를 사용자에게 보여주면서 작업을 위한 텍스트 노드를 관리하는 문서 트리 관리자(12); 일련의 마크업 문자열을, 프로그램이 접근할 수 있는 결과 구조로 변환하고, 상기 문서 트리 관리자(12)로부터의 마크업 처리 요구 및 유효성 또는 문법 검증 요구에 따라 마크업을 처리하면서 유효성이나 문법에 대한 검증을 수행하여 상기 문서 트리 관리자(12)로 보고하는 파서(14)(parser); 읽어 들인 XML 텍스트 문서 또는 비 XML 텍스트 문서를 대상으로 문서의 내용을 편집하는 XML 편집기(16); 현재 문서 변환 작업중인 문서상의 위치 또는 사용자가 선택한 노드에 따라서 삽입이 가능한 후보 엘리먼트를 제시해 주는 후보 엘리먼트 관리자(18); XML 구조정보를 적용하여 변환해야 할 텍스트에 대한 태깅이 가능한 후보 엘리먼트를 제시해 주는 텍스트 태깅 관리자(20); 문서상의 현재 위치하고 있는 엘리먼트(태그) 또는 선택한 엘리먼트가 속성을 가지고 있을 경우에 설정 가능한 속성 목록과 속성값을 제시해 주며 외부의 입력값에 따라 상기 속성값을 설정하거나 변경하는 속성 관리자(22); 및 읽어 들인 문서에서 에러를 체크하여 오류 사항을 표시해 주고, 오류 발생 항목에 커서를 이동시켜 주는 오류 관리자(24)를 구비한다.

상기 XML 편집기(16)는 상기 후보 엘리먼트 관리자(18)와 텍스트 태깅 관리자(20) 및 속성 관리자(22)로부터의 정보를 상기 문서 트리 관리자(12)를 통해 받아서 편집중인 문서에 구조 정보를 입력하고 그 속성을 변경한다.

또한, 상기 XML 편집기(16)는 사용자의 편의를 위해서 문서의 논리 구조에 해당하는 태그들과 내용에 해당하는 텍스트들에 대한 색상을 구별하여 설정하며, 텍스트 문서상에서도 문서의 논리 구조를 명확히 할 수 있도록 구조의 깊이에 따른 들여쓰기를 설정할 수도 있다.

상기 엘리먼트라 함은 XML 텍스트 문서의 기본이 되는 논리적 단위로서, XML 텍스트 문서의 모든 내용들은 반드시 엘리먼트 안에 포함되어야 한다. 다시 말해서, 상기 엘리먼트는 DTD(Document Type Definition : 문서 타입 정의)에서 태그와 논리적 구조를 선언하는 단위이다.

예를 들어, " < 슬라이드>

< 슬라이드> ..< /슬라이드>

< 슬라이드> ..< /슬라이드>

공개특허 등2002- 0043115

< /슬라이드쇼 > " 와 같은 XML 구조가 있다고 가정하였을 경우, 상기의 < 슬라이드쇼 > 엘리먼트는 두 개의 < 슬라이드 > 엘리먼트를 포함한다고 할 수 있다.

도 2는 본 발명의 방법이 적용되는 시스템의 작업 원도우의 예이다.

본 발명의 방법이 적용된 시스템의 작업 원도우는 현재 작업중인 문서(즉, XML 텍스트 문서, 비 XML 텍스트 문서)의 구조 정보를 표현하는 XML 트리 원도우(30); 현재 작업중인 문서를 텍스트 형태로 보여주는 XML 에디터 원도우(32); 작업중인 문서상의 위치 또는 사용자가 선택한 노드에 따라서 삽입이 가능한 후보 엘리먼트들을 제시해 주는 엘리먼트 후보 목록 원도우(34); 문서상의 현재 위치하고 있는 엘리먼트 또는 선택한 엘리먼트가 속성을 가지고 있을 경우에 설정 가능한 속성 목록과 속성값을 모두 보여주는 속성 목록 원도우(36); 작업중인 문서상의 XML 텍스트 또는 비 XML 텍스트의 일부영역을 선택하여 구조 정보를 적용시켜 XML 텍스트 문서로 변환시켜 나가기 위한 후보 엘리먼트를 제시해 주는 텍스트 태깅 후보 원도우(38); 읽어 들인 문서에서의 오류 사항을 표시해 주는 에러 리스트 원도우(40); 파일, 편집, 보기, 도구, 도움말 등과 같은 메뉴; 및 도구 바로 구성된다.

상기 XML 트리 원도우(30)는 사용자가 XML 텍스트 문서를 열기하였을 경우와 비 XML 텍스트 문서를 열기하였을 경우에 따라 상호 다른 형태의 화면을 제시한다.

즉, 사용자가 XML 텍스트 문서를 열기하였을 경우 상기 XML 트리 원도우(30)는 XML 텍스트 문서의 구조 정보를 보여준다.

그러나, 사용자가 비 XML 텍스트 문서를 열기하였을 경우 상기 XML 트리 원도우(30)는 문서 타입(DTD파일과 문서 타입)을 함께 지정하여 지정된 문서 타입을 근거로 최소화된 문서 구조를 보여준다. 이때, 비 XML 텍스트 영역은 XML로 변환하기 위한 작업대상 텍스트로 노드가 만들어진다.

상기 XML 트리 원도우(30)내의 트리에 있는 노드들을 대상으로 복사, 이동, 삭제 등의 작업을 할 수 있다.

그리고, 상기 XML 에디터 원도우(32)에서는 읽어 들인 XML 텍스트 문서 또는 비 XML 텍스트 문서를 대상으로 문서의 내용을 편집할 수도 있고, 상기 엘리먼트 후보 목록 원도우(34)와 속성 목록 원도우(36) 및 텍스트 태깅 후보 원도우(38)를 사용하여 편집중인 문서에 구조 정보를 입력하고 그 속성을 변경할 수 있다. 또한, XML 에디터 원도우(32)에서는 사용자의 편의를 위하여 문서의 논리 구조에 해당하는 태그들과 내용에 해당하는 텍스트들에 대한 색상을 구별하여 설정할 수 있으며, 텍스트 문서상에서도 문서의 논리 구조를 명확히 할 수 있도록 구조의 깊이에 따른 들어쓰기를 설정할 수도 있다.

상기 XML 트리 원도우(30)와 상기 XML 에디터 원도우(32)는 항상 동기화되어 있어서 트리의 노드를 선택하여 문서의 위치를 쉽게 찾아 볼 수 있게 한다.

문서의 특정 위치에서 삽입할 수 있는 엘리먼트들은 DTD(문서 타입 정의)에 의해 결정되는데, 상기 엘리먼트 후보 목록 원도우(34)에서는 엘리먼트를 삽입할 수 있는 경우에 문서상의 현재 위치 또는 작업을 하고자 하는 동작에 따라 후보 엘리먼트를 보여준다.

즉, 상기 엘리먼트 후보 목록 원도우(34)에는 다수의 세부 메뉴가 있는데, 선택되는 세부 메뉴에 따라 상기 원도우(34)에 디스플레이되는 후보 엘리먼트들이 다르게 나타난다.

즉, 상기 엘리먼트 후보 목록 원도우(34)는 현재 선택 또는 위치한 노드의 앞에 삽입할 수 있는 후보 엘리먼트를 제시해 주는 비포(before) 메뉴, 현재 선택 또는 위치한 노드의 뒤에 삽입할 수 있는 후보 엘리먼트를 제시해 주는 애프터

(after) 메뉴, 현재 선택 또는 위치한 노드의 마지막 자식으로 삽입할 수 있는 후보 엘리먼트를 제시해 주는 차일드(child) 메뉴, 현재 선택 또는 위치한 노드들의 부모로 삽입될 수 있는 후보 엘리먼트를 제시해 주는 서라운드(surround) 메뉴, 현재 선택 또는 위치한 노드의 엘리먼트 이름을 바꿀 수 있는 후보를 제시해 주되 현재 선택된 엘리먼트 자식 노드들을 가지고 있을 경우 자식 노드들은 그대로 유지하는 리네임(rename) 메뉴, 및 현재 선택 또는 위치한 노드들을 바꿀 수 있는 후보 엘리먼트를 제시해 주되 현재 선택된 노드들은 자식노드까지 모두 삭제되고 후보 엘리먼트들이 그 자리에 삽입되는 리플레이스(replace) 메뉴를 갖는다.

그리고, 상기 속성 목록 원도우(36)는 문서상의 현재 위치하고 있는 엘리먼트 또는 선택한 엘리먼트가 속성을 가지고 있을 경우에 설정 가능한 속성 목록과 속성값을 모두 보여주며 그 값을 설정하거나 변경할 수 있게 해준다. 여기서, 상기 속성에도 종류가 있다. 열거형인 속성은 설정할 수 있는 값이 이미 DTD에 정의되어 있다. 따라서, 열거형인 속성은 설정되어 있는 값들의 목록을 보여주어서 그 안에서만 선택할 수 있다. 다른 종류의 속성들은 사용자 입력으로 변경할 수 있다.

상기 XML 트리 원도우(30) 및 XML 에디터 원도우(32)는 상기 문서 트리 관리자(12) 및 그와 연동하는 XML 편집기(16)의 동작에 의해 출력되는 결과들을 표시해 주는 창이고, 상기 엘리먼트 후보 목록 원도우(34)는 상기 문서 트리 관리자(12) 및 그와 연동하는 후보 엘리먼트 관리자(18)의 동작에 의해 출력되는 결과들을 표시해 주는 창이며, 상기 속성 목록 원도우(36)는 상기 문서 트리 관리자(12) 및 그와 연동하는 속성 관리자(22)의 동작에 의해 출력되는 결과들을 표시해 주는 창이고, 상기 텍스트 태깅 후보 원도우(38)는 상기 문서 트리 관리자(10) 및 그와 연동하는 텍스트 태깅 관리자(20)의 동작에 의해 출력되는 결과들을 표시해 주는 창이며, 상기 에러 리스트 원도우(40)는 상기 오류 관리자(24)의 동작에 의해 출력되는 결과들을 표시해 주는 창이다.

도 3은 본 발명의 방법이 적용된 시스템에서 문서 열기하여 문서 편집과정으로 진입하기까지의 동작을 설명하는 흐름도이다.

먼저, 사용자가 도 2의 초기화면에서 파일 메뉴를 클릭한 후 찾고자 하는 파일이 있는 위치에서 파일 열기를 할 파일의 이름 및 그 파일의 형식(XML 파일 또는 텍스트 파일)을 지정한 후 해당 파일을 열게 된다.

즉, 상기 사용자가 XML 파일을 선택한 경우(단계 S10에서 "XML 파일") 그 XML 파일 선택신호는 XML 편집기(16)를 통해 문서 트리 관리자(12)를 거쳐 파서(14)로 입력되고, 그 파서(14)에서 해당 XML 파일을 파싱하여 문서의 구조 정보 및 데이터를 추출(단계 S12)한 후 그 결과를 상기 문서 트리 관리자(12)로 보낸다.

그에 따라, 상기 문서 트리 관리자(12)는 DOM+(10)과 연동하여 XML 편집기(16)를 통해 도 4에 예시한 바와 같이 XML 트리 원도우(30)에 상기 XML 파일을 파싱한 구조 정보로 DOM트리를 구성하고, XML 에디터 원도우(32)에 상기 읽어들인 XML 파일(즉, 구조화된 XML 텍스트 파일)을 나타낸다(단계 S14).

이에 반해, 상기 사용자가 텍스트 파일을 선택한 경우(단계 S10에서 "텍스트 파일") XML 파일을 템플릿으로 사용하는지의 여부에 따라 이후의 동작이 달라진다.

상기 사용자가 텍스트 파일을 선택하고 XML 파일을 템플릿으로 사용하는 경우 즉, 기존의 XML 문서로 변환시킨 것을 이용하여 텍스트 파일을 XML 문서로 변환하고 싶어 할 때(단계 S16에서 "Yes")에는 상기 텍스트 파일 선택신호 및 XML 파일의 템플릿 사용신호는 상기 XML 편집기(16)를 통해 문서 트리 관리자(12)를 거쳐 파서(14)로 입력되고, 그 파서(14)에서 해당 XML 파일을 파싱하여 문서의 구조 정보만 추출(단계 S18)하여 그 결과를 상기 문서 트리 관리자(12)로 보낸다. 상기 문서 트리 관리자(12)는 상기 XML 파일내의 텍스트 부분은 모두 삭제한다.

그에 따라, 상기 문서 트리 관리자(12)는 DOM+(10)과 연동하여 XML 편집기(16)를 통해 도 5에 예시한 바와 같이 XML 트리 원도우(30)에 상기 XML 파일을 파싱한 구조 정보로 DOM트리를 구성하고 그 DOM트리에 텍스트 파일 내용의 텍스트 노드를 추가하며, 상기 XML 에디터 원도우(32)에는 상기 읽어들인 텍스트 내용(즉, 구조화되지 않은 비 XML 텍스트 영역이 존재하는 텍스트)을 나타낸다(단계 S20).

한편, 상기 사용자가 텍스트 파일을 선택하고 XML 파일을 템플릿으로 사용하지 않는 경우 즉, 선택한 텍스트 파일만을 이용하여 문서 편집을 하고 싶어 할 경우(단계 S16에서 "No") 상기 사용자는 도 6에 예시한 바와 같이 DTD 파일 및 DOC TYPE을 지정한다(단계 S22). 상기 사용자가 미리 DTD 파일 및 DOC TYPE을 지정해 둘 수도 있다. 이와 같이 DTD 파일 및 DOC TYPE이 미리 지정되어 있는 경우에는 확인 버튼만 클릭하면 된다.

이후, 상기 문서 트리 관리자(12)는 파서(14) 및 DOM+(10)과 연동하여 XML 편집기(16)를 통해 도 7에 예시한 바와 같이 XML 트리 원도우(30)에 DTD에 근거한 최소한의 DOM트리를 구성하고 그 DOM트리에 텍스트 내용의 텍스트 노드를 추가하며, 상기 XML 에디터 원도우(32)에는 상기 읽어들인 텍스트 내용을 나타낸다(단계 S24).

도 8은 본 발명의 방법이 적용된 시스템에서 XML 문서 열기에서의 오류 처리과정을 설명하는 흐름도이다.

사용자가 도 2의 초기화면에서 파일 메뉴를 클릭한 후 찾고자 하는 파일이 있는 위치에서 파일 열기를 할 파일의 이름 및 그 파일의 형식(XML 파일)을 지정한 후 해당 파일을 선택한다.

그에 따라, XML 편집기(16)는 파서(14)를 통하여 선택된 XML 문서를 파싱하여 문서상에 오류가 있는지를 검증한다(단계 S30).

그 XML 문서에 오류가 발견되지 않은 경우(단계 S32에서 "No")에는 그 파서(14)는 문서의 구조정보 및 데이터를 추출하여 문서 트리 관리자(12)에게 보내고 상술한 XML 문서열기와 같은 과정을 거치게 된다(단계 S34).

반면에, 상기 XML 문서에 오류가 있는 경우(단계 S32에서 "Yes")에는 그 파서(14)는 최대한 오류에 대한 처리를 진행하면서 에러 리스트를 작성하여 오류 관리자(24)에게 전달하고 그 오류 관리자(24)는 도 9에 예시한 바와 같이 에러 리스트 원도우(40)상에 소정의 에러 리스트를 표시한다. 그 오류 관리자(24)는 오류 메시지 디스플레이와 함께 오류가 발생된 위치에 커서를 자동으로 이동시켜 줄 수도 있고, 오류 메시지가 다수개이면 어느 한 오류 메시지를 선택하면 해당 오류 위치로 커서가 자동으로 이동하게 할 수도 있다.

또한, 상기 오류 관리자(24)는 전달받은 에러 리스트를 근거로 하여 XML 문서에 부적합한 텍스트 부분이 있는지를 판단한다(단계 S36).

부적합한 텍스트 부분이 없는 경우(단계 S36에서 "No") 즉, 작업대상인 비 XML 텍스트가 있는 것이 아니라 예를 들어 엘리먼트의 속성이 누락되어 있거나 시작 태그는 있는데 끝 태그가 없는 등의 경우 상기 오류 관리자(24)는 오류에 대한 처리(즉, 누락된 속성을 제 위치에 삽입해 주거나 누락된 끝 태그를 삽입해 주는 일 등)를 진행하면서 예러 처리 항목(예컨대, 어느 엘리먼트의 속성이 누락되어 삽입하였습니다. 어느 엘리먼트의 끝 태그가 누락되어 삽입하였습니다. 등과 같은 예러 처리 항목)을 작성하여 도 9에 예시한 바와 같이 에러 리스트 원도우(40)상에 표시한다(단계 S38, S40).

상기 판단 결과, 상기 단계 S36에서 부적합한 텍스트가 있는 경우 즉, 작업대상 텍스트(비 XML 텍스트)가 존재한 상태로 작업을 종료하고 저장하였을 경우 단계 S42로 이동하여 작업대상 텍스트로 인식하여 문서 트리 관리자(12)는 도 5 또는 도 7에서와 같이 XML 트리 원도우(30)에 상기 XML 파일을 파싱한 구조 정보로 DOM트리를 구성하고 그 DOM트리에 부적합한 텍스트 내용의 텍스트 노드를 추가하며, 상기 XML 에디터 원도우(32)에는 XML 텍스트 영역 및 비

XML 텍스트 영역이 존재하는 XML 문서를 나타낸다. 여기서, 상기 비 XML 텍스트 영역은 아직 구조 정보가 적용되지 않은 텍스트 영역으로서, 작업자는 이 영역에 대하여 태깅을 하여 구조화된 XML 텍스트 문서로 완성시킨다.

도 10은 본 발명의 실시예에서 텍스트 태깅 후보 엘리먼트를 제시하는 과정을 설명하는 흐름도이다.

사용자가 구조정보를 적용하여 XML 문서로 변환하고자 하는 일반 텍스트 문서 또는 변환 작업이 완료되지 않은 XML 문서를 지정하면 XML 편집기(16)는 해당 파일을 읽어서 구조 정보는 XML 트리 윈도우(30)에 디스플레이시키고, 텍스트 파일 내용은 XML 에디터 윈도우(32)상에 디스플레이시킨다.

그 상태에서, 사용자가 구조정보를 적용하여 XML로 변환시킬 비 XML 텍스트 또는 추가적으로 구조정보를 적용시킬 XML 텍스트에서 임의의 영역을 선택하면(단계 S50에서 " Yes") 예를 들어도 11에서와 같이 구조 정보를 적용시켜야 할 비 XML 텍스트 영역(31)중에서 " 기술연구소" 를 선택하게 되면 그에 상응하는 선택신호가 텍스트 태깅 관리자(20)로 입력된다.

그에 따라, 상기 텍스트 태깅 관리자(20)는 선택된 영역(" 기술연구소")의 부모 노드에 삽입가능한 엘리먼트가 존재하는지를 판단하는데, 이 경우 선택된 영역(" 기술연구소")의 앞 및 뒤에 있는 노드들을 같이 비교하여 부모 노드에 삽입할 수 있는 후보 엘리먼트가 있는지를 판단한다(단계 S52).

상기 판단결과, 상기 선택된 영역(" 기술연구소")의 부모 노드에 삽입가능한 엘리먼트가 존재한다고 판단되면(단계 S52에서 " Yes") 상기 텍스트 태깅 관리자(20)는 텍스트 태깅 후보 윈도우(38)에 태깅할 수 있는 후보 엘리먼트(예컨대, " employee")를 제시한다(단계 S54).

그리고, 상기 텍스트 태깅 관리자(20)는 사용자로부터 엘리먼트 태깅 선택 또는 엘리먼트 하부 선택이 있는지를 판단한다(단계 S56).

사용자가 상기 엘리먼트 태깅을 선택하게 되면(단계 S56에서 " 태깅") 그에 상응하는 선택신호가 텍스트 태깅 관리자(20)로 입력되고, 상기 텍스트 태깅 관리자(20)는 상기 선택된 영역(" 기술연구소")을 상기 선택한 후보 엘리먼트의 내용으로 태그가 삽입된다(단계 S58).

반면에, 사용자가 상기 엘리먼트 하부를 선택하게 되면(단계 S56에서 " 하부") 그에 상응하는 선택신호가 텍스트 태깅 관리자(20)로 입력되고, 상기 텍스트 태깅 관리자(20)는 " employee" 자식으로 나올 수 있는 후보 엘리먼트(예컨대, " name" , " role" , " email" 등)를 파악하여 도 12와 같이 텍스트 태깅 후보 윈도우(38)상에 제시한다(단계 S60). 여기서, 상기 " employee" 자식중 " role" 를 선택하면 " role" 자식으로 나올 수 있는 후보 엘리먼트(" dept" , " team" , " jikgeup" , " jikcheak")를 제시한다. 그 중 " dept" 를 선택하면 자식으로 나올 후보가 없으므로 더 이상 자식 후보가 표시되지 않는다.

상기 사용자가 텍스트 태깅 후보 윈도우(38)상의 " dept" 를 더블 클릭하게 되면 그에 상응하는 더블 클릭신호가 텍스트 태깅 관리자(20)로 입력되고, 상기 텍스트 태깅 관리자(20)는 도 13에서와 같이 XML 에디터 윈도우(32)상에 " 기술연구소" 에 태그를 적용시킨다. 상기 태깅될 때 문서 구조 정보(DTD 파일)에 맞게 다른 엘리먼트들이 생성되고, 나머지 텍스트들은 작업대상 텍스트 노드로 남게 된다.

그리고, 상기 XML 에디터 윈도우(32)와 상기 XML 트리 윈도우(30)는 항상 동기화되기 때문에 상기 XML 트리 윈도우(30)에도 역시 그에 상응하는 DOM 트리가 구성된다.

한편, 본 발명의 실시예는 작업대상인 텍스트를 처음부터 차례대로 순차적으로 태깅할 수 있을 뿐만 아니라, 랜덤하게 자신이 희망하는 텍스트 영역을 선택하여 텍스트 태깅 후보 윈도우(38)상에 디스플레이되는 후보 엘리먼트를 보고서 태깅할 수도 있다.

발명의 효과

이상 상세히 설명한 바와 같이 본 발명에 따르면, 읽어들인 XML 텍스트 문서 또는 비 XML 텍스트 문서를 대상으로 상황에 맞는 후보 엘리먼트를 제시하여 선택되는 후보 엘리먼트의 내용으로 태그를 삽입시킴으로써, XML 텍스트 문서로의 변환시 간편하게 태깅할 수 있게 해 주고, 그로 인해 텍스트 내용에 구조 정보를 적용시키는 과정 및 시간이 단축되어 종래에 비해 효율적이다.

한편, 본 발명은 상술한 실시예로만 한정되는 것이 아니라 본 발명의 요지를 벗어나지 않는 범위내에서 수정 및 변형하여 실시할 수 있고, 그러한 수정 및 변형이 가해진 기술사상 역시 이하의 특허청구범위에 속하는 것으로 보아야 한다.

(57) 청구의 범위

청구항 1.

변환하기 위한 비 XML 텍스트 및 변환 과정에서 생성된 XML 텍스트를 하나의 문서로 통합하여 관리하는 제 1과정; 및

상기 XML 텍스트 및 상기 비 XML 텍스트의 선택된 영역에 대해 XML 문서 규칙에 따른 적용가능한 태깅정보를 제공하여 논리 구조를 적용시키는 제 2과정을 포함하는 것을 특징으로 하는 텍스트 태깅 방법.

청구항 2.

편집 영역에 나타나는 문서내의 XML 텍스트 영역 및 비 XML 텍스트 영역중에서 일부 영역이 선택되면 상기 선택된 일부 영역에 태깅가능한 후보 엘리먼트의 존재여부를 파악하여 화면상에 표시하는 제 1과정; 및

상기 화면표시된 후보 엘리먼트에 대한 태깅이 선택되면 상기 선택된 일부 영역에 상기 선택된 엘리먼트의 내용으로 태그를 삽입하고, 상기 표시된 후보 엘리먼트에 대한 하부 엘리먼트 선택이 있으면 상기 선택된 후보 엘리먼트의 하부 엘리먼트를 조사하여 화면상에 표시하여 태깅 선택을 대기하는 제 2과정을 포함하는 것을 특징으로 하는 텍스트 태깅 방법.

청구항 3.

제 2항에 있어서,

상기 제 1과정은 상기 선택된 일부 영역의 전 및 후에 있는 노드들을 같이 비교하여 부모 노드에 삽입할 수 있는 후보 엘리먼트를 결정하여 화면상에 표시해 주는 것을 특징으로 하는 텍스트 태깅 방법.

청구항 4.

제 2항에 있어서,

상기 화면상에 디스플레이되는 후보 엘리먼트 및 하부 엘리먼트는 상기 화면의 소정 영역에 위치한 텍스트 태깅 후보 원도우상에 디스플레이되는 것을 특징으로 하는 텍스트 태깅 방법.

청구항 5.

제 2항에 있어서,

상기 일부 영역은 작업자의 임의대로 선택 가능한 것을 특징으로 하는 텍스트 태깅 방법.

청구항 6.

제 2항에 있어서,

상기 편집 영역에 나타나는 비 XML 텍스트 영역에 대해서는 상기 화면의 소정 영역에 위치한 트리 윈도우상에 그 비 XML 텍스트 영역의 노드를 표시하는 것을 특징으로 하는 텍스트 태깅 방법.

청구항 7.

변환하기 위한 비 XML 텍스트 및 변환 과정에서 생성된 XML 텍스트를 하나의 문서로 통합하여 관리하는 제 1과정; 및

상기 XML 텍스트 및 상기 비 XML 텍스트의 선택된 영역에 대해 XML 문서 규칙에 따른 적용가능한 태깅정보를 제공하여 논리 구조를 적용시키는 제 2과정의 순서로 진행되는 텍스트 태깅 프로그램이 저장된 컴퓨터에서 재생가능한 기록매체.

청구항 8.

편집 영역에 나타나는 문서내의 XML 텍스트 영역 및 비 XML 텍스트 영역중에서 일부 영역이 선택되면 상기 선택된 일부 영역에 태깅가능한 후보 엘리먼트의 존재여부를 파악하여 화면상에 표시하는 제 1과정; 및

상기 화면표시된 후보 엘리먼트에 대한 태깅이 선택되면 상기 선택된 일부 영역에 상기 선택된 엘리먼트의 내용으로 태그를 삽입하고, 상기 표시된 후보 엘리먼트에 대한 하부 엘리먼트 선택이 있으면 상기 선택된 후보 엘리먼트의 하부 엘리먼트를 조사하여 화면상에 표시하여 태깅 선택을 대기하는 제 2과정의 순서로 진행되는 텍스트 태깅 프로그램이 저장된 컴퓨터에서 재생가능한 기록매체.

청구항 9.

제 8항에 있어서,

상기 제 1과정은 상기 선택된 일부 영역의 전 및 후에 있는 노드들을 같이 비교하여 부모 노드에 삽입할 수 있는 후보 엘리먼트를 결정하여 화면상에 표시해 주는 것을 특징으로 하는 컴퓨터에서 재생가능한 기록매체.

청구항 10.

제 8항에 있어서,

상기 화면상에 디스플레이되는 후보 엘리먼트 및 하부 엘리먼트는 상기 화면의 소정 영역에 위치한 텍스트 태깅 후보 윈도우상에 디스플레이되는 것을 특징으로 하는 컴퓨터에서 재생가능한 기록매체.

청구항 11.

제 8항에 있어서,

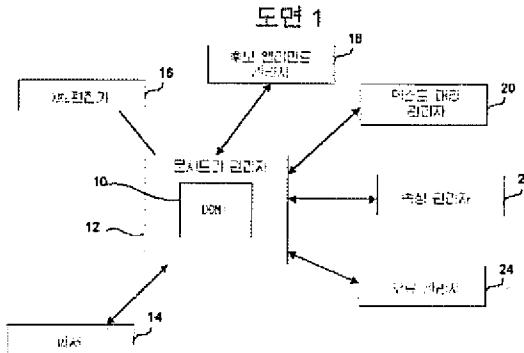
상기 일부 영역은 작업자의 임의대로 선택 가능한 것을 특징으로 하는 컴퓨터에서 재생가능한 기록매체.

청구항 12.

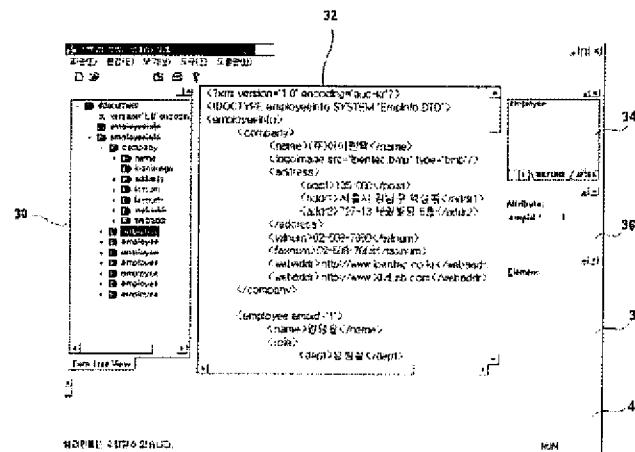
제 8항에 있어서,

상기 편집 영역에 나타나는 비 XML 텍스트 영역에 대해서는 상기 화면의 소정 영역에 위치한 트리 윈도우상에 그 비 XML 텍스트 영역의 노드를 표시하는 것을 특징으로 하는 컴퓨터에서 재생 가능한 기록매체.

도연

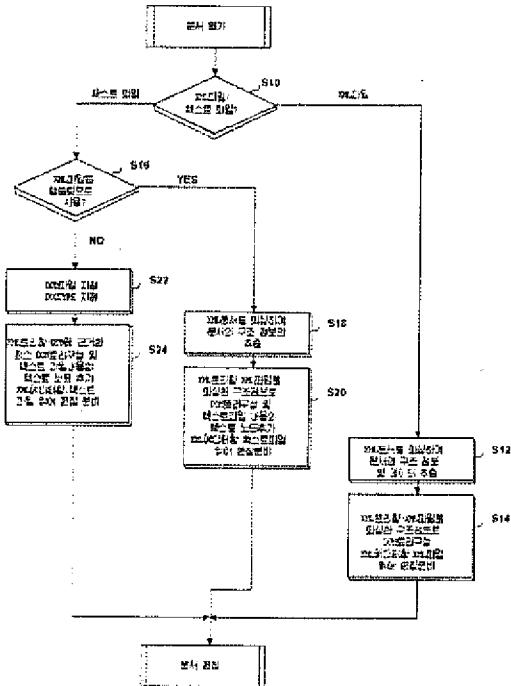


도면 2

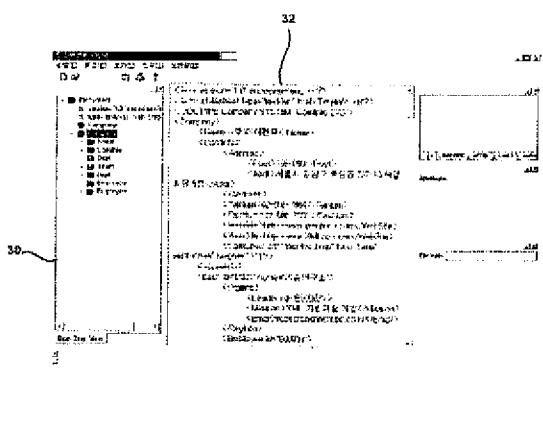


공개특허 제2002- 0043115

도면 3

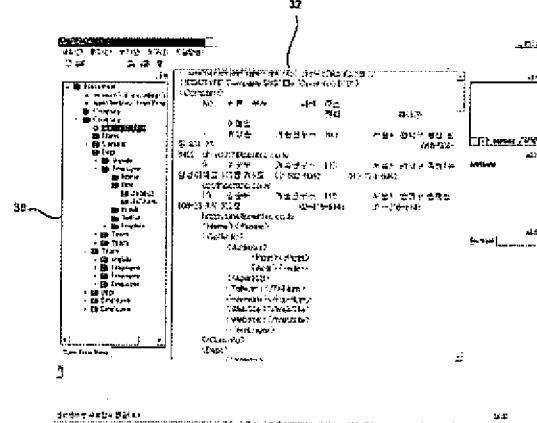


도면 4

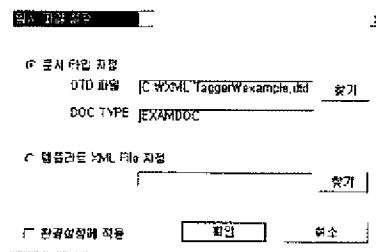


광개특허 특2002- 0043115

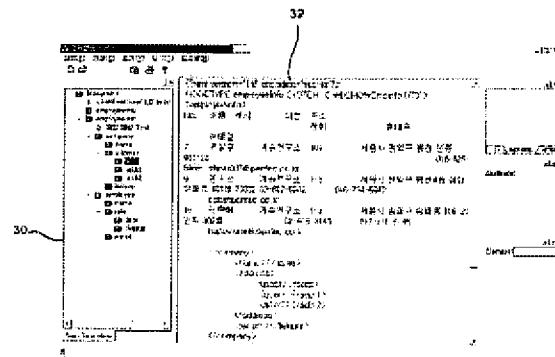
도면 5



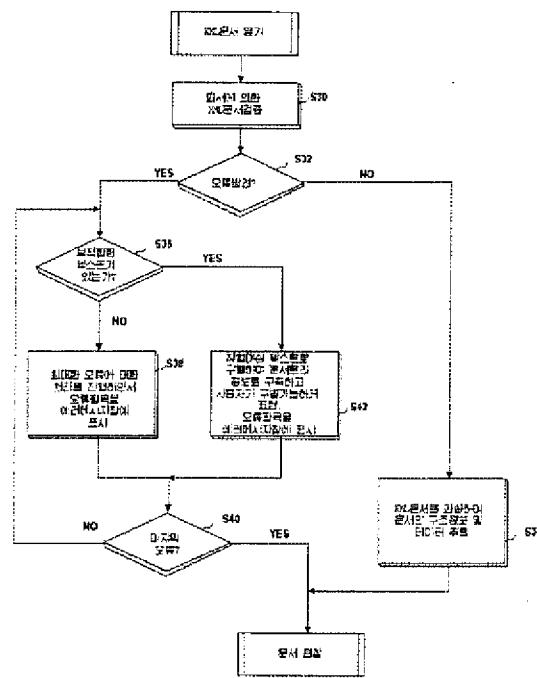
五四 6



도면 7

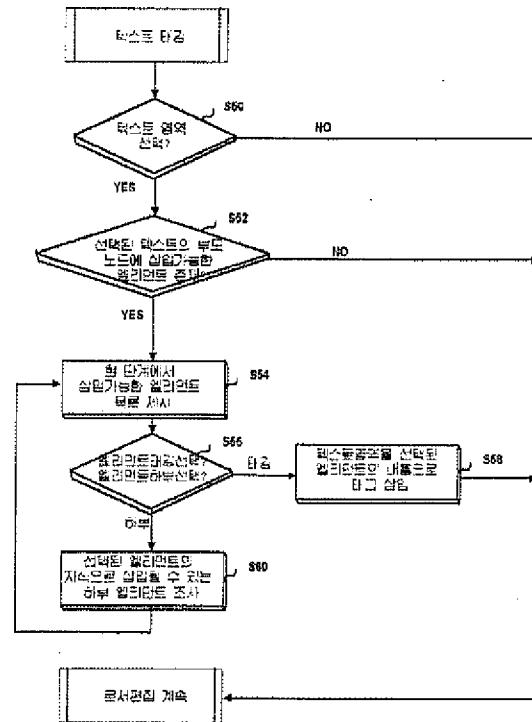


도면 8

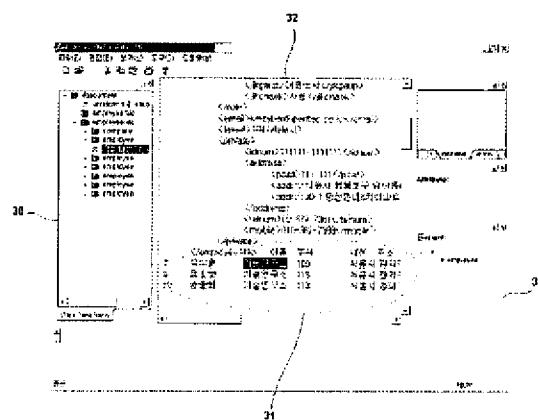


도면 9

도면 10

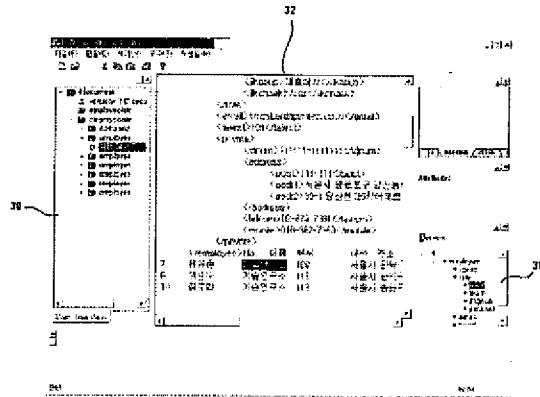


도면 11



공개특허 특2002- 0043115

도면 12



도면 13

